

FIGURE 1

(SEQ ID NO:1-nucleotide sequence and SEQ ID NO:2-amino acid sequence)
Escherichia coli B Phytase Sequence

1
ATG AAA GCG ATC TTA ATC CCA TTT TTA TCT CTT CTG ATT CCG TTA ACC CCG
Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr Pro

CAA TCT GCA TTC GCT CAG AGT GAG CCG GAG CTG AAG CTG GAA AGT GTG GTG
Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser Val Val

ATT GTC AGT CGT CAT GGT GTG CGT GCT CCA ACC AAG GCC ACG CAA CTG ATG
Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr Gln Leu Met

CAG GAT GTC ACC CCA GAC GCA TGG CCA ACC TGG CCG GTA AAA CTG GGT TGG
Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val Lys Leu Gly Trp

CTG ACA CCG CGN GGT GGT GAG CTA ATC GCC TAT CTC GGA CAT TAC CAA CGC
Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu Gly His Tyr Gln Arg

CAG CGT CTG GTA GCC GAC GGA TTG CTG GCG AAA AAG GGC TGC CCG CAG TCT
Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys Lys Gly Cys Pro Gln Ser

GGT CAG GTC GCG ATT ATT GCT GAT GTC GAC GAG CGT ACC CGT AAA ACA GGC
Gly Gln Val Ala Ile Ile Ala Asp Val Asp Glu Arg Thr Arg Lys Thr Gly

GAA GCC TTC GCC GCC GGG CTG GCA CCT GAC TGT GCA ATA ACC GTA CAT ACC
Glu Ala Phe Ala Ala Gly Leu Ala Pro Asp Cys Ala Ile Thr Val His Thr

CAG GCA GAT ACG TCC AGT CCC GAT CCG TTA TTT AAT CCT CTA AAA ACT GGC
Gln Ala Asp Thr Ser Ser Pro Asp Pro Leu Phe Asn Pro Leu Lys Thr Gly

GTT TGC CAA CTG GAT AAC GCG AAC GTG ACT GAC GCG ATC CTC AGC AGG GCA
Val Cys Gln Leu Asp Asn Ala Asn Val Thr Asp Ala Ile Leu Ser Arg Ala

GGA GGG TCA ATT GCT GAC TTT ACC GGG CAT CGG CAA ACG GCG TTT CGC GAA
Gly Gly Ser Ile Ala Asp Phe Thr Gly His Arg Gln Thr Ala Phe Arg Glu

CTG GAA CGG GTG CTT AAT TTT CCG CAA TCA AAC TTG TGC CTT AAA CGT GAG
Leu Glu Arg Val Leu Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu

AAA CAG GAC GAA AGC TGT TCA TTA ACG CAG GCA TTA CCA TCG GAA CTC AAG
Lys Gln Asp Glu Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys

GTG AGC GCC GAC AAT GTC TCA TTA ACC GGT GCG GTA AGC CTC GCA TCA ATG
Val Ser Ala Asp Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met

CTG ACG GAG ATA TTT CTC CTG CAA CAA GCA CAG GGA ATG CCG GAG CCG GGG
Leu Thr Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly

TGG GGA AGG ATC ACC GAT TCA CAC CAG TGG AAC ACC TTG CTA AGT TTG CAT
Trp Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His

AAC GCG CAA TTT TAT TTG CTA CAA CGC ACG CCA GAG GTT GCC CGC AGC CGC
Asn Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser Arg

GCC ACC CCG TTA TTG GAT TTG ATC ATG GCA GCG TTG ACG CCC CAT CCA CCG
Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His Pro Pro

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CAA	AAA	CAG	GCG	TAT	GGT	GTG	ACA	TTA	CCC	ACT	TCA	GTA	CTG	TTT	ATT	GCC
Gln	Lys	Gln	Ala	Tyr	Gly	Val	Thr	Leu	Pro	Thr	Ser	Val	Leu	Phe	Ile	Ala
GGA	CAC	GAT	ACT	AAT	CTG	GCA	AAT	CTC	GGC	GGC	GCA	CTG	GAG	CTC	AAC	TGG
Gly	His	Asp	Thr	Asn	Leu	Ala	Asn	Leu	Gly	Gly	Ala	Leu	Glu	Leu	Asn	Trp
ACG	CTT	CCC	GGT	CAG	CCG	GAT	AAC	ACG	CCG	CCA	GGT	GGT	GAA	CTG	GTG	TTT
Thr	Leu	Pro	Gly	Gln	Pro	Asp	Asn	Thr	Pro	Pro	Gly	Gly	Glu	Leu	Val	Phe
GAA	CGC	TGG	CGT	CGG	CTA	AGC	GAT	AAC	AGC	CAG	TGG	ATT	CAG	GTT	TCG	CTG
Glu	Arg	Trp	Arg	Arg	Leu	Ser	Asp	Asn	Ser	Gln	Trp	Ile	Gln	Val	Ser	Leu
GTC	TTC	CAG	ACT	TTA	CAG	CAG	ATG	CGT	GAT	AAA	ACG	CCG	CTG	TCA	TTA	AAT
Val	Phe	Gln	Thr	Leu	Gln	Gln	Met	Arg	Asp	Lys	Thr	Pro	Leu	Ser	Leu	Asn
ACG	CCG	CCC	GGA	GAG	GTG	AAA	CTG	ACC	CTG	GCA	GGA	TGT	GAA	GAG	CGA	AAT
Thr	Pro	Pro	Gly	Glu	Val	Lys	Leu	Thr	Leu	Ala	Gly	Cys	Glu	Glu	Arg	Asn
GCG	CAG	GGC	ATG	TGT	TCG	TTG	GCA	GGT	TTT	ACG	CAA	ATC	GTG	AAT	GAA	GCA
Ala	Gln	Gly	Met	Cys	Ser	Leu	Ala	Gly	Phe	Thr	Gln	Ile	Val	Asn	Glu	Ala
CGC	ATA	CCG	GCG	TGC	AGT	TTG	AGA	TCT	CAT	CAC	CAT	CAC	CAT	CAC	TAA	1323
Arg	Ile	Pro	Ala	Cys	Ser	Leu	Arg	Ser	His	His	His	His	His	His	End	

[illegible]

FIGURE 2
pH/Temperature Profile and Stability

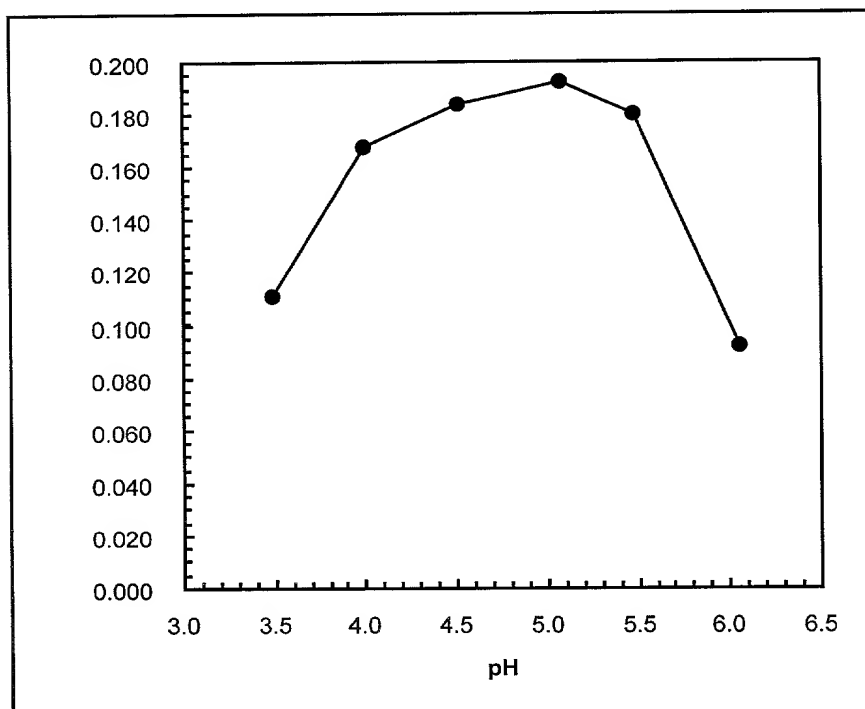
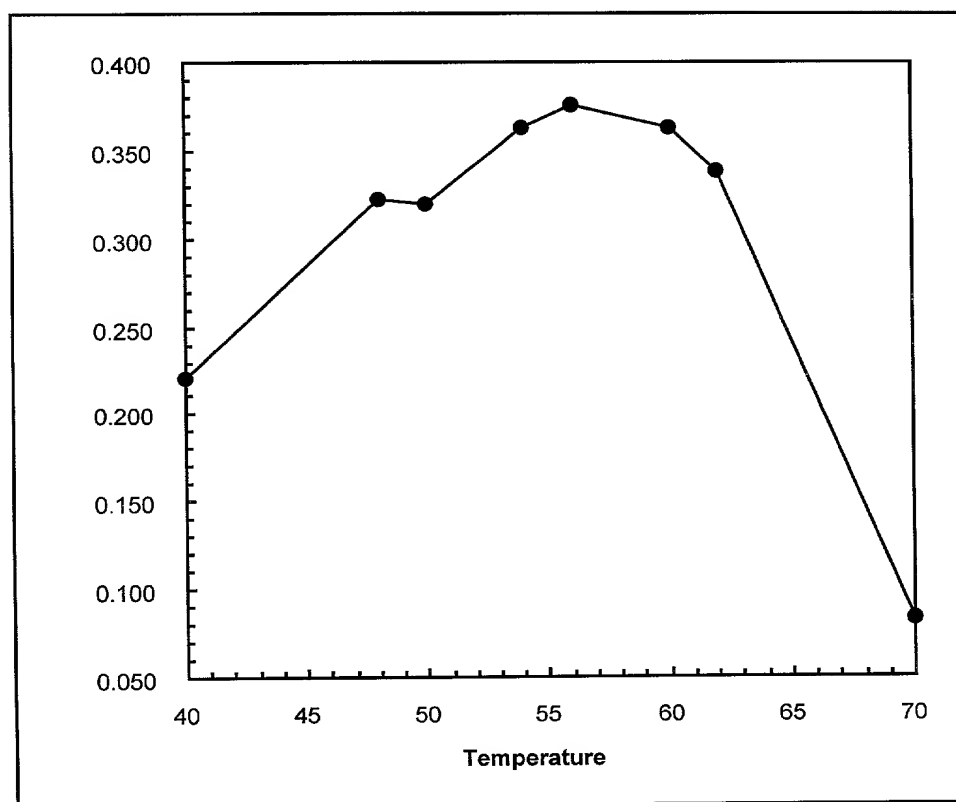


Figure 3:

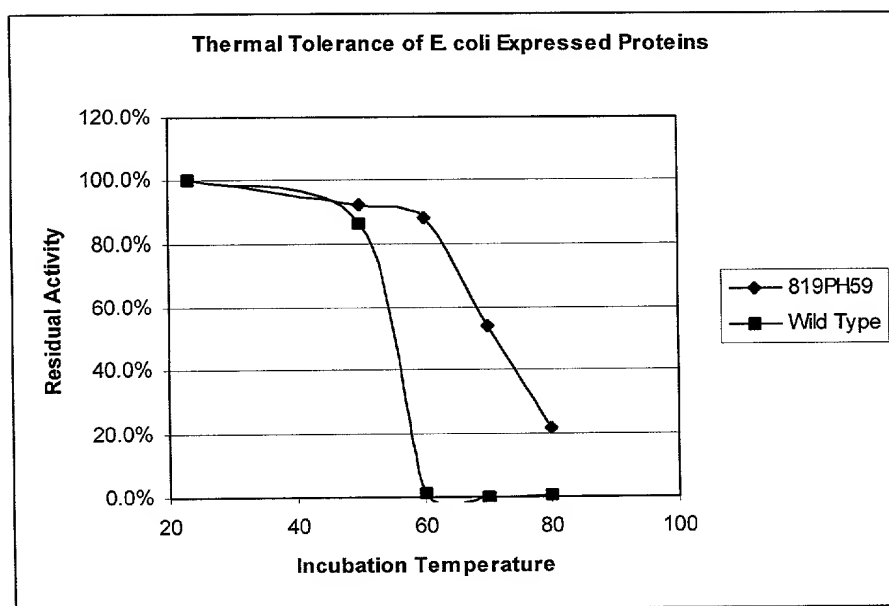


Figure 4

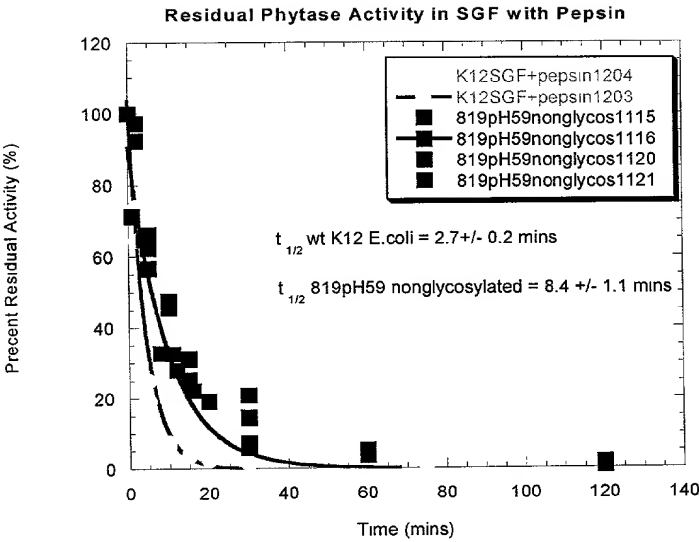


Figure 5:

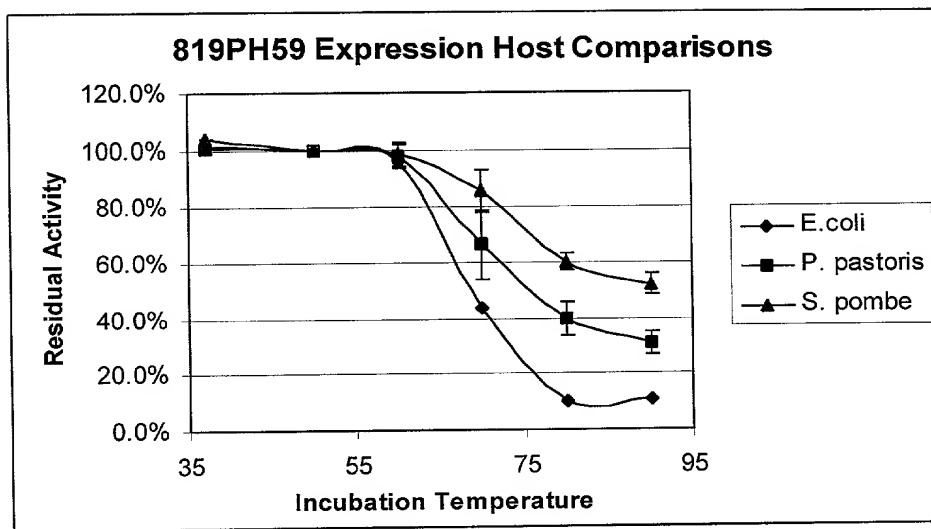


Figure 6.

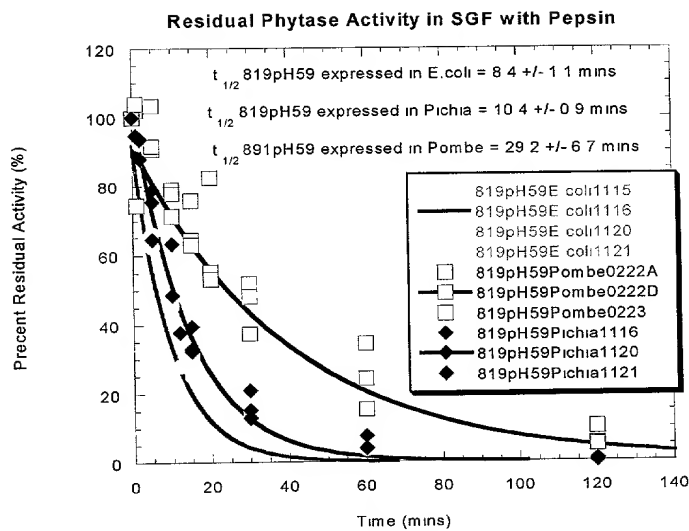


Figure 7a

E. coli appA (GenBank accession no. M58708) (SEQ ID NO:7)

1 taaggagcag aaacaatgtg gtatttactt tggttcgtcg gcattttgtt gatgtgttcg
 61 ctctccaccc ttgtgttggt atggctggac ccgcgtctga aaagttaacg aacgtaggcc
 121 tgatcgccgc cattagcatc gcatcaggca atcaataatg tcagatatga aaagcggaaa
 181 catatcgatg aaagcgatct taatcccatt ttatctctt ctgattccgt taaccccgca
 241 atctgcattc gtcagagtg agccggagct gaagctggaa agtgtggtga ttgtcagtcg
 301 tcatgggtg cgtgctccaa ccaaggccac gcaactgatg caggatgtca ccccagacgc
 361 atggccaacc tggccggtaa aactgggttg gctgacaccg cgnngtggtg agctaategc
 421 ctatctcgga cattaccaac gccagcgtct ggtagccgac ggattgctgg cgaanaaggg
 481 ctgcccgcag tctggtcagg tcgcattat tctgatgtc gacgagcgt cccgtaaaa
 541 aggcgaagcc ttcgcccgcg ggctggcacc tgactgtgca ataaccgtac ataccaggc
 601 agatacgtcc agtcccgtac cgttatttaa tctctaaaa actggcggtt gccaaactgga
 661 taacgcgaac gtgactgacg cgatcctcag cagggcagga gggtaattg ctgacttac
 721 cgggcatcgg caaacggcgt ttcgcgaact ggaacgggtg cttaatttc cgcaatcaa
 781 ctgtgcctt aaacgtgaga aacaggacga aagctgttca ttaacgcagg cattaccatc
 841 ggaactcaag gtgagcgccg acaatgtctc attaacgggt gcgtaagcc tcgcatcaat
 901 gctgacggag atatttctcc tgcaacaagc acagggaatg ccggagccgg ggtggggaag
 961 gatcccgat tcacaccagt ggaacacctt gctaagttg cataacgcgc aattttatt
 1021 gctacaacgc acgccagagg ttgcccgag ccgcgccacc ccgttattag atttgatcaa
 1081 gacagcgttg acgccccatc caccgcaaaa acaggcgtat ggtgtgacat taccacttc
 1141 agtgtgttt atcgccggac acgatactaa tctggcaaat ctggcgccg cactggagct
 1201 caactggacg ctccccggtc agccggataa cacgccgcca ggtggtgaac ttgtgttga
 1261 acgctggcgt cggtaagcg ataacagcca gtggattcag gtttcgtgg tcttcagac
 1321 ttacagcag atgcgtgata aaacgccgct gtcattaaat acgccgccc gagaggtgaa
 1381 actgaccctg gcaggatgtg aagagcgaaa tgcgcagggc atgtgtcgt tggcaggtt
 1441 tacgcaaate gtgaatgaag cagcatacc ggcgtgcagt ttgtaatgca taaaaagag
 1501 cattcagtta cctgaatgct ctgaggctga tgacaaacga agaactgtct aatgcgtaga
 1561 ccggaanaag cgttcaagcc gcatccggcc acitfcagtt ttctcttc tcggagtaac
 1621 tataaccgta atagttag ccgtaactgt aagcgggtgt ggcgcgtta atcacaccat
 1681 tgaggatagc gccttaata ttgacgctg cctgttcag acgtgcatt gacaaactca
 1741 cctcttggc ggtgttcaag ccaaacgcg caaccagcag gctggtgcca acagaacgcc
 1801 ccacgacgc ggcatcactc accgccagca tcggcgccgt atcgacaatc accagatcgt
 1861 aatggtcgtt cgccattcc agtaattgac gcatccgac g

T04250"6459960

Figure 7b

1 taaggagcag aaacaatgtg gtatttactt tggttcgtcg gcattttgtt gatgtgttcg
 61 ctctccaccc ttgtgttggt atggctggac ccgcgtctga aaagttaacg aacgtaggcc
 121 tgatcgggcg cattagcatc gcatcaggca atcaataatg tcagatatga aaagcggaaa
 181 catatcgatg aaagcgatct taatcccatt tttatctctt ctgattccgt taaccccgca
 241 atctgcattc gctcagagtg agccggagct gaagctggaa agtgtggtga ttgtcagtcg
 301 tcatggtgtg cgtgctccaa ccaaggccac gcaactgatg caggatgtca cccagacgc
 361 atggccaacc tggccggtaa aactgggttg gctgacaccg cngggtggtg agctaattcg
 421 ctatctcgga cattaccaac gccagcgtct ggtagccgac ggattgctgg cgaaaaaggg
 481 ctgcccgcag tctggtcagg tcgcgattat tgcgatgtc gacgagcgtc cccgtaaaac
 541 aggccaagcc ttcgccgccg ggctggcacc tgactgtgca ataaccgtac ataccaggc
 601 agatacgtcc agtcccgtac cgttatttaa tctctaaaa actggcgttt gccactgga
 661 taacgcgaac gtgactgacg cgatcctcag cagggcagga gggtaattg ctgacttac
 721 cgggcatcgg caaacggcgt ttcggaact ggaacgggtg cttatattc cgcaatcaa
 781 cttgtgcctt aaacgtgaga aacaggacga aagctgtca ttaacgcagg cattaccatc
 841 ggaactcaag gtgagcgccg acaatgtctc attaacgggt gcggtgaagc tcgcatcaat
 901 gctgacggag atatttctcc tgcaacaagc acagggaatg ccggagccgg ggtggggaag
 961 gatcaccgat tcacaccagt ggaacacctt gctaagttg cataacgcgc aattttattt
 1021 gctacaacgc acgccagagg ttgccgcag ccgcgccacc ccgttattag atttgatcaa
 1081 gacagcgttg acgccccatc caccgcaaaa acaggcgtat ggtgtgacat taccacttc
 1141 agtgcgtgtt atcgccggac acgatactaa tctggcaaat ctggcgggcg cactggagct
 1201 caactggacg ctccccggtc agccggataa cacgccgcca ggtggtgaac ttgtgtttga
 1261 acgctggcgt cggctaagcg ataacagcca gtggattcag gtttcgctgg tctccagac
 1321 ttacagcag atgcgtgata aaacgccgct gtcattaaat acgccgccc gagaggtgaa
 1381 actgaccctg gcaggatgtg aagagcgaaa tgcgcagggc atgtgttcgt tggcaggttt
 1441 tacgcaaata gtgaatgaag cacgcatacc ggctgcagt ttgtaatgca taaaaagag
 1501 cattcagtta cctgaatgct ctgaggctga tgacaaacga agaactgtct aatgcgtaga
 1561 ccggaagagg cgttcacgcc gcacccggcc actttcagtt ttctctttc tcggagtaac
 1621 tataaccgta atagttatag ccgtaactgt aagcgggtct ggcgcgttta atcacacat
 1681 tgaggatagc gcctttaata ttgacgcctg cctgttcag acgctgcatt gacaaactca
 1741 cctctttggc ggtgttcaag ccaaacgcg caaccagcag gctggtgcca acagaacgcc
 1801 ccacgaccgc ggcatactc accgccagca tcggcggcgt atcgacaatc accagatcgt
 1861 aatggtcgtt cgccattcc agtaattgac gcacccgac g

DIVER1370-7

Figure 8

Amino acid sequence for E. coli appA (wild type) (SEQ ID NO:8)

MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVT
 PDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQV
 AIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLD
 NA
 NVTDAILSRAGGSIADFTGHRRQTAFRELERVLNFPQSNLCLKREKQDESCSLT
 QALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQW
 NTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPSTV
 LFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQ
 VSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVN
 EARIPACSL

Bold-Underlined amino acid residues are shown below in the modified appA enzyme
 (SEQ ID NO:10)

MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVT
 PDAWPTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVA
 IIAADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDN
 A
 NVTDAILEEAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQ
 ALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWN
 TLLSLHNAQFDLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPSTVLF
 IAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVS
 LVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEA
 RIPACSL

T04250" 6429880